Nicole R. Robertson. The Perceptions of Traumatic Brain Injury Researchers Toward Data Sharing. A Master's Paper for the M.S. in I.S. degree. November 2019. 44 pages. Advisor: Helen Tibbo

Data curators have encouraged researchers to share their study data with repositories because new knowledge often comes from the secondary analysis of verifiable data sets. Data sets that are shared with repositories are also protected from loss and technological obsolescence, so their content may be fully exploited over time. However, many members of the traumatic brain injury (TBI) research community are not convinced that data sharing is good policy. This exploratory study seeks to identify the reasons why these scientists are or are not in favor of sharing study data with repositories. This study also seeks to identify which of the prevailing reasons in favor of sharing data these scientists are aware of and if they agree with them. In conclusion, this study suggests more phenomena that may be worth examining: the costs associated with sharing TBI study data, the data submission requirements of repositories, and the reuse of shared TBI study data.

Headings:

Brain injuries – Research

Information sharing

Intellectual cooperation

Information technology – Scientific applications

THE PERCEPTIONS OF TRAUMATIC BRAIN INJURY RESEARCHERS TOWARD DATA SHARING

by
Nicole R. Robertson

A Master's paper submitted to the faculty
of the School of Information and Library Science
of the University of North Carolina at Chapel Hill
in partial fulfillment of the requirements
for the degree of Master of Science in
Information Science.

Chapel Hill, North Carolina

November 2019

Approved by

_____

Helen Tibbo

Acknowledgements

First, I wish to thank those members of my Team Mental Health: all of my family but especially my loving and inspiring parents, Angeline and James; my heroically patient significant other and live-in tech support, Jared; my ever-helpful friends Paul, Kim, and Sue. I also wish to thank those members of my Team Physical Health, because without their guidance I wouldn't have remained fit enough to complete this degree, let alone the work you are about to read. They include Dr. Thomas Koinis, Dr. Anne Toohey, and Yonica Rodio, DPT. Next, I wish to thank Denise Anthony and Helen Tibbo of the School of Information and Library Science for all the times they treated me less like a student and more like a colleague. Finally, I wish to thank those members of the traumatic brain injury research community who offered me their encouragement as I progressed through the data collection phase of my study.

It really does take a village to conduct research!

Table of Contents

Introduction

In her 2010 feature article titled “The Imperative for Data Curation,” Joyce L. Ogburn described the data sharing policies of the National Institutes of Health and the National Science Foundation as “public investments in science” that are “predicated on the idea that the sharing of research data and publishable results stimulates additional innovation and discoveries” (p. 241). Ogburn also postulated that this “open system of knowledge demands an infrastructure that will endure well into the future. Leaving digitally based information to languish in personal electronic filing drawers amid a jumble of unrelated information and with no plans for its survival guarantees its disappearance” (pp. 241-242).

Given this warning about the dangers of condemning research data to death via technological obsolescence and neglect, one might believe that all scientists naturally support the concept of data sharing. Yet this is not necessarily the case, as Felicia LeClere pointed out in an opinion piece published later that same year: “Data sharing is a bit like going to the dentist. We can all agree that it is a good thing to do and intrinsic to good scientific practice. In reality, however, researchers tend to view data sharing with a mix of fear, contempt, and dread.”

Although I have not had as many opportunities to discuss the merits and disadvantages of data sharing as LeClere, I am aware that some scientists would prefer not to upload their data sets to a repository where others may access them. In May 2017, I completed a set of interviews and a Data Curation Profile (http://datacurationprofiles.org/overview.php) on the Assessment of Military Multitasking Performance project, which ran from August 2012 through August 2015.

The instrument generated by this study targets known sensorimotor, cognitive, and physical vulnerabilities in patients who have sustained a mild traumatic brain injury and is intended to assist with return-to-duty decision-making regarding military service members who have suffered such injuries (Weightman et al., 2017). During the second interview, a co-investigator associated with the project surprised me by saying, "I guess I feel like this whole issue of data sharing is something that is just becoming a 'thing.' So I know moving forward, when you're funded it's got to be something that you plan for but at the time this project was funded, it wasn't really a requirement" (N. Robertson, Personal interview, March 27, 2017). When I listened to these comments again as I typed the transcript of that interview, I wondered if this researcher was aware of the reasons why data ought to be shared, since the researcher seemed to believe that data management plans were simply the latest trend in grant application requirements. Now I wonder if other scientists engaged in traumatic brain injury research are making the same assumption or if their opinions reflect the mix of fear, contempt, and dread described by LeClere. Or do their opinions reveal a completely different relationship with data sharing than the ones LeClere and I have put forward?

Given my interest in these researchers and their perceptions of data sharing via repositories, this exploratory study will focus on the following questions:

- Are these scientists reluctant to preserve and share their research data with repositories and if so, why?
- Are these scientists aware of the prevailing reasons for sharing research data? Do the scientists who are aware of these reasons agree with them?
- How did these scientists form their opinions about sharing research data?

Literature Review

The case for sharing research data through repositories

Ever since the advent of cyberscholarship – which is broadly defined as the move from laboratory-based science to science based on large quantities of data that are analyzed by computers – data curators have encouraged researchers to share the data sets they produce with the wider scientific community and even the public. Perhaps the oldest and most compelling reason given by data curators is that new knowledge often comes from the secondary analysis of verifiable data sets that are made available to other researchers (Oliver & Harvey, 2016, p. 22). Yet there are other good reasons for scientists to share their data sets, as well.

Data sets placed with repositories are protected from loss and technological obsolescence, so their content may be fully exploited by researchers over time. Repositories will also provide information about the context and provenance of the data sets for this same reason, which in turn maximizes the investment made by the original researcher who produced the data sets in the first place. In addition, the results of research paid for with public funds must often be shared without restriction, so the accompanying data sets must be placed in repositories that will make them available to the public (Oliver & Harvey, 2016, pp. 11-13).

In some cases, scientists are legally required to manage and retain their data sets for certain research projects, which makes them prime candidates for deposit into a repository (Digital Curation Centre, 2017). For example, the United Kingdom's Natural Environment Research Council must comply with Antarctic Treaty regulations; that is,

all data collected from the Antarctic (south of 60 degrees) must be disseminated beyond members of the Council (Oliver & Harvey, 2016, p. 13).

In other cases, scientists must manage and retain their data sets because they are required to do so by the funder of their research project. The United States National Science Foundation (NSF) decided to take this step in 2010, when it announced that all proposals for the 2011 grant cycle had to include a plan for sharing "the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants" with other researchers (National Science Foundation, 2017, Chapter XI). The United States Department of Defense (DoD) followed in the NSF's footsteps a few years later when it published its *Plan to Establish Public Access to the Results of Federally Funded Research* in 2015. According to this plan, a supplementary data management document was to become an integral part of all contract or grant proposal packages by the beginning of the 2016 fiscal year. The DoD stated that such documents would likely need to include "the types of data, software, curriculum materials, and other materials to be produced in the course of the project that are publicly releasable" (Department of Defense, 2015, p. 3). In addition, these documents would need to include the conditions for access, sharing, re-use, re-distribution, and creation of derivative works based on the submitted data.

The case for sharing data on traumatic brain injury research through repositories

Scientists who study patients with traumatic brain injuries (TBIs) also have good reasons to share their data through repositories, although very little is written about them. However, some reasons for data sharing can be found in the policy documentation of the repositories themselves; for example, the data sharing policy of the Federal Interagency Traumatic Brain Injury Research Informatics System (FITBIR) states, "The

goal of the policy is to advance science for the benefit of the public through the creation of a centralized Federal data repository for TBI research information" (FITBIR, 2014, p. 3). Along similar lines, the research collaboration policy of the TBI Endpoints Development (TED) Initiative, which manages the TED Metadataset, states, "Access to study data, materials sharing, and mutual collaboration among research teams in order to accelerate research in TBI are fundamental tenets of the TED project and are core beliefs of its investigators" (TED Initiative, 2015, p. 3). These sentiments of data sharing for the public good and to accelerate TBI research are echoed in the external use notification of the TBI Model Systems National Database, which is maintained by the TBI Model Systems National Data and Statistical Center. It states, "The TBI Model Systems Centers Program has a responsibility to the public in general, and to the scientific community in particular, to encourage scientific use of the TBI Model Systems National Database and archived Collaborative Module Study Databases as rapidly as possible, subject to appropriate terms and conditions" (TBI Model Systems National Data and Statistical Center, 2015, p. 1).

In August 2012, however, the TBI research community received a directive from the highest-ranking officer of the United States government concerning its data sharing practices. In that month, President Barack Obama issued Executive Order 13625, titled "Improving Access to Mental Health Services for Veterans, Service Members, and Military Families." The order, which sought to "build an integrated network of support capable of providing effective mental health services for veterans, service members, and their families," also recognized that the "lack of full understanding of the underlying mechanisms of Post-Traumatic Stress Disorder (PTSD), other mental health conditions, and Traumatic Brain Injury (TBI) has hampered progress in prevention, diagnosis, and treatment" (Section 1 and Section 5). Bearing this latter point in mind, the Order also decreed that the Departments of Defense, Veterans Affairs, Health and Human Services,

and Education should establish a National Research Action Plan. In addition to enhancing the medical community's understanding of the mechanisms responsible for PTSD, related injuries, and neurological disorders following TBI, the National Research Action Plan was to "improve data sharing between agencies and academic and industry researchers to accelerate progress and reduce redundant efforts without compromising privacy" (Exec. Order No. 13625, 2012, Section 5). This led the Department of Veterans Affairs to report that they were "sharing data as appropriate" with the four agencies mentioned above, as well as colleges and universities. The report also announced that Veterans Affairs was establishing new data-sharing agreements, although the report does not list which agencies or institutions are involved (U.S. Department of Veterans Affairs, 2017).

The medical research community's attitudes toward data sharing

In spite of all this, many members of the medical research community are still not convinced that data sharing is good policy. For example, the United Kingdom's Expert Advisory Group on Data Access reported that clinical research data would only be shared if there were more incentives for scientists to do so. One such incentive included adequate levels of recognition for scientists who chose to share their data, either in project funding, career advancement, or recognition with the United Kingdom's Research Excellence Framework (O'Dowd, 2014). Members of the group also stated that research institutions should provide training and support for researchers to manage data effectively, as those in some disciplines did not have access to the skills and resources they needed to share their data easily. In light of this particular recommendation, the group proposed that the career pathways for data managers should be strengthened, so that scientists would have access to such skilled personnel (O'Dowd, 2014).

A study on the data-sharing attitudes and practices of the scientific and clinical research staffs served by the U.S. National Institutes of Health Library also revealed that nearly half of the survey respondents never had an opportunity to share their data with a repository (Federer, Lu, Joubert, Welsh, & Brandys, 2015). Other reasons the survey respondents gave for not sharing data included having no knowledge of any repositories that accepted the types of data they produced, fear of compromising their research subjects' privacy because their data included personally identifying information, believing that it is not customary to share data in their research fields, and not knowing how to prepare their documentation and data for sharing with others.

Medical researchers have also found fault with the proposed data-sharing policies of medical journals, as in 2016 when two groups of clinical investigators opposed a plan set forth by the International Committee of Medical Journal Editors (ICMJE). According to the plan, any researcher who wished to register a randomized, controlled clinical trial and publish the results in an ICMJE journal had to agree to share the de-identified patient data underlying the study within six months of publication. A spokesperson for one of the groups, called the International Consortium of Investigators for Fairness in Trial Data Sharing, responded by saying, "Is it really fair that we should be the ones footing all those bills and then simply handing the data over to someone else? Or is it not appropriate that the people who were involved in the study have the first opportunity to publish on the data that was generated from the studies that we conducted and based upon the money that we helped to invest to get those trials done?" (Murray, 2016, E336)

This same consortium published a more detailed position statement on the website of the New England Journal of Medicine a few months after the ICMJE outlined its plan. Much of this statement expressed the group's dissatisfaction with the ICMJE's time frame for data release and proposed instead that it allow investigators exclusive use of their data for a minimum of 2 years after the primary trial results were published. If

the trial had taken more than 2 years to complete, then an additional 6 months should be added to this figure for every year it took to complete the trial, up to a maximum of 5 years. The consortium believed these limits would give the original investigators ample time to prepare secondary analyses of their data, which they claimed was a "key motivation" for researchers to conduct and publish the results of clinical trials in the first place. In addition, the consortium felt that it was only fair to give the original investigators first opportunity to perform these new analyses because they had expended so much effort to finish the trial and produce the data (International Consortium of Investigators for Fairness in Trial Data Sharing, 2016).

Along with this alternate time frame for data release, the consortium also suggested that investigators ought to be compensated financially for access to the data generated by their trials. They argued that if the ICMJE's plan was enacted and all clinical trialists were suddenly required to share their data as a condition of publication, then the preparation of data for release "would probably divert resources, both financial and human, from the actual conduct of trials" (International Consortium of Investigators for Fairness in Trial Data Sharing, 2016). The consortium also stated that many clinical trialists spend a great deal of money generated from other activities to fund trials that are not of interest to commercial sponsors. Therefore, the consortium felt that investigators should be able to charge anyone requesting access to their data, in order to recoup some of the costs of those trials ignored by commercial sponsors and to "fund the data-preparation activities necessary for data sharing in such a way as to protect confidentiality and ensure data integrity" (International Consortium of Investigators for Fairness in Trial Data Sharing, 2016).

While the consortium debated the perceived merits and disadvantages of the ICMJE's data sharing plan, a task force within the Coordinated Research Infrastructures Building Enduring Life-Science Services Project (CORBEL) produced ten principles and

fifty associated recommendations that represented what it saw as “the fundamental requirements for any framework for the sharing and reuse of clinical trials data” (Ohmann et al., 2017, p. 4). This task force, which consisted primarily of medical researchers, patient representatives, funders, information technologists, and other experts from the European Union, developed these principles and recommendations by means of a consensus-building procedure called the nominal group process. During this process, a core group first proposed a text to the larger group and then a moderator collected comments from each group member and collapsed the similar ones. Next, the core group prioritized a series of discussion points based on these comments and all group members discussed and voted on each one. Then the core group revised the text based on the voting results and the whole process began again until consensus was reached. By using these procedures, the CORBEL task force reached a consensus on their principles and recommendations over the course of one year, from March 2016 to March 2017 (Ohmann et al., 2017).

Among the principles agreed upon by the task force is one that endorses obtaining explicit broad consent from trial participants for the sharing and reuse of their data for scientific purposes only. Another favors de-identifying data sets to protect the privacy of trial participants, as well as recording any de-identification steps that are applied. Still another suggests limiting access to only those citizens or groups that have both a reasonable scientific question and the expertise to answer that question based on the shared data. Yet the two principles that set the CORBEL task force apart from the consortium are the ones that fully support the sharing of data and documents (such as study protocols and case report forms) from clinical trials and the placement of such information with suitable repositories. Although the members of the task force came to a consensus regarding these principles, they recognized that some other investigators still have concerns about “the resources required to support data preparation, or potential

misinterpretation of their data, or a possible reduction in the number of papers they will be able to generate from the data themselves" (Ohmann et al., 2017, pp. 4-5). However, the task force concluded that these concerns could be mitigated if changes were made to the way clinical research is recognized and rewarded and if appropriate resources, policies, and procedures were established to make data sharing a normal part of clinical research (Ohmann et al., 2017).

Based on this summation of the literature, it is evident that many members of the medical research community are still reluctant to participate in data sharing; however, most of these objections are not new. In a 2010 opinion piece, research and data scientist Felicia LeClere describes several eerily familiar reasons other scientists have given her over the years for declining to share their own data sets. First and foremost of these reasons is the desire to protect the identities of research subjects - a valid concern that was echoed by the clinical and scientific research staffs served by the U.S. National Institutes of Health Library and the Coordinated Research Infrastructures Building Enduring Life-Science Services Project. Another is the desire to exploit one's data as much as possible via secondary analyses because the original investigator worked hard to obtain the data, which was echoed by the International Consortium of Investigators for Fairness in Trial Data Sharing. Still another is the belief that data sets require substantial amounts of time and effort to clean up for other scientists to use, which was echoed by both the International Consortium of Investigators for Fairness in Trial Data Sharing and the United Kingdom's Expert Advisory Group on Data Access. Although data curators may be reassured by the fact that they have the talking points necessary to address these objections with potential users of their repositories, they should be concerned that many scientists have resisted or never encountered the arguments made in favor of data sharing over the years.

The TBI research community's attitudes toward data sharing

Just like their colleagues in the larger medical research community, scientists who study patients with traumatic brain injuries (TBIs) also have opinions about sharing research data with others in their field. Unfortunately, these opinions are harder to discover because little has been written about them. The information that exists comes from the International Initiative for Traumatic Brain Injury Research (InTBIR), an organization that describes itself as "a cooperative effort of the European Commission, the Canadian Institutes of Health Research, and the National Institutes of Health to coordinate and leverage clinical research activities on traumatic brain injury research" (International Initiative for Traumatic Brain Injury Research, 2014, InTBIR Mission). Because the organization wishes to promote data sharing across the field of TBI research and is currently working to implement a data sharing policy, in 2016 it surveyed the content of information sheets and informed consent forms given to patients with TBIs in seven out of eleven InTBIR projects. As stated in the final report, the aim of the survey was "to determine whether and to what extent data sharing is possible with the content currently adopted by the InTBIR consortia, and to propose common content based on adaptation or extension of currently used wording" (International Initiative for Traumatic Brain Injury Research, 2016b, Purpose section). As a result, the organization found that the majority of consortia did not make any specific provisions for the transfer of patient data between different research centers or countries, although consent to do so was implied in many cases. From this, it is possible to infer that TBI researchers may see the lack of informed patient consent as a barrier to the sharing of their research data. However, the organization also determined that the lack of such provisions was merely an oversight on the part of the consortia and not due to any deliberate policies. For these reasons, the surveyors concluded that InTBIR project teams should clearly explain any arrangements for the sharing of data within the information sheets or informed consent

forms they provide to patients (International Initiative for Traumatic Brain Injury Research, 2016b).

In October of the same year, all those who attended InTBIR's fifth annual meeting in Washington, DC received another survey from the leaders of the meeting's final session, titled "Priority Objectives and Milestones for 2017." The survey's purpose was to identify high priority aims and objectives for the following year, along with the action items and workgroups needed to achieve them. (The leaders of the final session intended to collect this information during the meeting but time constraints and a limited number of participants prevented them from doing so.) In addition to gathering suggestions for such aims and objectives, the survey also asked the attendees to list action items that would help to achieve the data sharing needed for international collaboration. Although some of the answers to this question varied widely – from the familiar call for the de-identification of patient data to a request for a list of repositories "qualified" to host TBI data – most of them proved that the attendees were well aware of the technological issues that had to be resolved before data sharing could begin. For example, one attendee recommended that InTBIR metadata should be uploaded into accessible databases maintained by entities such as the Ontario Brain Institute, the Human Brain Project, and the Federal Interagency Traumatic Brain Injury Research Informatics System; once completed, a federated search function should then be provided so that researchers could identify relevant data sets for collaborative projects. Another attendee proposed that "state of the science" tools for data analysis and visualization should be provided to InTBIR investigators along with any shared data sets. Still others advocated for the provision of a searchable database of biospecimens and a simple broker to manage the exchange of InTBIR data between different countries (International Initiative for Traumatic Brain Injury Research, 2016a). As with the previous survey on the content of the information sheets and informed consent forms

given to patients with TBIs, it is possible to infer that TBI researchers may see the absence of these technologies as additional barriers to the sharing of their research data.

Those who attended InTBIR's first annual meeting naturally discussed this global approach to the sharing of TBI data at some length, as evidenced by the conference report published in the Journal of Neurotrauma. During the session on the topic of "Bottlenecks and Priorities for Action," for example, the participants determined that the intellectual property rights of pharmaceutical companies should be protected in order to encourage their involvement within InTBIR. Therefore, the participants recommended that the organization find ways to strike a balance between the risks and benefits of sharing data on drug development (Tosetti et al., 2013). During the session on the topic of "Collecting and Sharing TBI Patient Data," however, many more concerns emerged. Several of these included methods for ensuring long-term access to patient data, determining who will have access to that data, methods for validating any data that is shared, and developing a system of governance for such a data sharing system (Tosetti et al., 2013). Here there is no need for inference; these TBI investigators clearly view the absence of controls over intellectual property rights, data that cannot be sustained long-term, ambiguous access policies, lack of data validation, and poor system governance as further barriers to the sharing of their research data.

As useful as this information may be for decoding the attitudes of TBI researchers towards data sharing, however, a quick reading of the introduction to the conference proceedings reveals that these opinions came from the "60 policy makers, scientists, clinicians, and patient and industry representatives" who attended from the United States, Canada, China, Australia, and European Union (Tosetti et al., 2013, p. 1211). In addition, the proceedings fail to mention how many of those attendees actually conduct TBI research, as opposed to those who represent the interests of patients or make decisions about funding policies.

The two surveys mentioned previously suffer from even smaller response rates than the sessions of InTBIR's first conference: the first survey received only seven out of eleven possible responses while the second received only five responses. When one remembers that these few sources represent what is currently known about what TBI researchers think of data sharing, it is natural to wonder if a greater variety of opinions exist within this population. Therefore, this exploratory study seeks to better understand these scientists' perceptions of data sharing and their willingness to share their own data by surveying a larger and more representative sample of this population.

Methodology

Survey construction and administration

A survey was created and distributed to members of the TBI research community to collect information on their perceptions of research data sharing and their willingness to share data via repositories. In order to protect the privacy of those who participated in the study, the survey did not ask for any personally identifying information, such as names or email addresses, and all responses to the survey were recorded anonymously. The survey contained a total of 16 closed and open-ended questions, although participants were presented with only 13 or 14 of these depending on their answer to the survey's single branching question. A complete list of questions included in the survey may be found in Appendix B.

The survey covered such topics as:

- Whether scientists studying traumatic brain injuries are in favor or not in favor of sharing research data via repositories and the reasons why they hold those beliefs;
- How scientists studying traumatic brain injuries formed their opinions about sharing research data (e.g., from conversations with their colleagues, administrators, and/or data curators, from publications or conference sessions about data sharing initiatives, etc.);
- Whether scientists studying traumatic brain injuries are aware of the prevailing reasons for sharing research data and if they agree with those reasons;

- Whether those scientists who are in favor of data sharing have ever deposited their research data with a repository and why they chose that particular repository;
- Whether those scientists who are not in favor of data sharing believe it is possible to resolve their concerns and if they would support data sharing under such circumstances, and;
- Whether scientists studying traumatic brain injuries have ever been required to deposit research data from one of their studies with a repository but had no such requirement for another study and how they would compare the two experiences.

The Qualtrics survey management platform was used to administer a Web-based version of the survey to the sample population. This method was chosen over paper and email surveys, phone interviews, focus groups, and individual interviews for several reasons. First, members of the TBI research community either work with or are familiar with several advanced technologies (such as the positron emission tomography and computed tomography scanners frequently used in neuroimaging), so it was likely that they would be comfortable working with Web-based surveys. Second, many members of the sample population resided in countries outside the United States, which would have made the administration of other types of survey instruments costly and inefficient. Third, a Web-based survey would help to eliminate researcher bias, especially when compared to phone interviews, focus groups, and individual interviews. As someone who wishes to enter the field of data curation and is a supporter of data sharing, I would have found it difficult to use neutral language if I had administered the survey over the phone or in person. Finally, the Qualtrics platform allowed respondents to save their progress and return to the survey at a later time if necessary; this feature typically helps to boost

response rates as respondents are able to complete the survey at their own pace (Wildemuth, 2009, pp. 257, 260).

Participant recruitment

The survey was first distributed to a cross-section of investigators who work within the TBI research community. This included academics and clinicians, as well as those who focus on the injuries of children, military service members, collegiate and professional athletes. The survey also targeted investigators who conduct their research both inside and outside the United States, since traumatic brain injuries are a leading cause of death and disability worldwide and are studied internationally (Allen Institute for Brain Science, 2014). For this reason, the sample population included 157 researchers from the United States and 93 researchers from countries such as Sweden, Canada, Finland, Australia, and the United Kingdom.

The sample population was chosen via snowball sampling. To select a snowball sample, one first identifies a few eligible members of the sample and then asks each of these participants for suggestions for additional people who meet the inclusion and exclusion criteria (Wildemuth, 2009, p. 121). Researchers who published studies in the *Journal of Neurotrauma* during the years 2019 through 2016 served as the first members of the snowball sample (n = 250) and were asked to forward the initial recruitment and follow-up emails to additional researchers who might agree to be surveyed. (The complete text of the initial recruitment email may be found in Appendix A.) The *Journal of Neurotrauma* was chosen for this purpose because it is the highly-regarded, official journal of the National Neurotrauma Society and the International Neurotrauma Society. It also regularly publishes studies on all types of traumatic brain injuries and lists the contact information of at least one participating investigator for each study it publishes.

Results and Discussion

Frequency of data sharing, expectations of others, and forming opinions

By the end of the two-week recruitment period, the Qualtrics survey management platform had collected a total of 38 responses. Unfortunately, nine of these were not usable as they only contained answers to the first question, the first three questions, or no questions at all. Once these incomplete responses were withdrawn, 29 remained available for analysis. As all survey responses were recorded anonymously, it is not known how many came from the original members of the snowball sample and how many came from TBI researchers who received a link to the survey from the members of this group.

The survey's first question asked participants how often traumatic brain injury researchers should share their study data with repositories. Sixteen participants answered "most of the time," seven participants answered "all of the time," five participants answered "some of the time," and one participant answered "seldom." None of the participants answered "never."

When asked to briefly explain their answer to this question, two participants who chose "most of the time" stated that "some data may be proprietary" and "some parts of data may not be permissible by local governing bodies for sharing." In another such case, the participant expressed the desire to share data from their own studies but lamented that their employer did not allow them to do so most of the time. In yet another such case, the participant voiced the opinion that TBI researchers who collect mathematically complex data sets should describe what is needed for other researchers to use them, to

avoid the "useless duplication" of uploading such data sets into repositories without the benefit of these insights from the original researchers. (It is worth noting that not all of the participants explained their answer to this question.)

Two participants who answered "some of the time" also gave their opinions on the policies that repositories ought to follow before accepting TBI study data for deposit. The first stated that small studies do not add value to a repository's holdings given the variance between all studies that is compounded when multiple data sets are merged. Therefore, this participant believed that repositories should apply "a rigor and scope threshold" before determining which data sets to accept. The second stated that data sharing depended on several factors, including how repositories controlled access to their holdings and the procedures they required researchers to follow in order to obtain data sets. As for the lone participant who answered "seldom," this researcher stated that submitting TBI study data to repositories is a very costly and time-consuming process for which no funding is provided. "In theory it's fine," the researcher wrote, "but practically it's not possible."

The survey's next question asked participants whether they would be concerned if certain people or entities expected them to share their study data with repositories. Participants were allowed to select all the responses that applied and a space was provided for the addition of responses that were not included in the question. "The organization that provided me with a research grant" was selected once, "my colleagues" was selected twice, and "my employer" was selected three times. "A publisher that accepted one of my manuscripts" was selected eight times and the most frequent response, "this expectation would not concern me," was selected sixteen times. It is worth noting that the most frequent response was the only one chosen in answer to this question by all but one of these sixteen participants. This particular researcher is also the one who disclosed that their employer limits how much data they can share and

indicated that they would be concerned if "a publisher that accepted one of my manuscripts" expected them to share their study data with repositories. When asked to briefly explain their answer to this question, the researcher wrote that such an expectation from a publisher "would be problematic for me, as I am not permitted to do so (even though I want to.)"

When the other participants were asked to briefly explain their answer to this question, some of their responses mirrored those opinions given by the individual researchers, consortiums, groups, and task forces mentioned in this study's review of the literature. (Again, readers should be aware that not all of the participants explained their answer.) For example, one researcher wrote that the expectation of data sharing did not concern them as long as the data was "exclusively available for a period of time to myself and my immediate collaborators." Determining just how long this period of time should last was the crux of the argument between the International Committee of Medical Journal Editors and the International Consortium of Investigators for Fairness in Trial Data Sharing. Another researcher wrote that they would be concerned if any person or entity expected them to share their study data with repositories "and did not provide funds for the effort it takes to process the data for submission." This sentiment was also expressed by the members of the United Kingdom's Expert Advisory Group on Data Access, who believed that they should be provided with the resources needed to share their data easily.

The survey's next question asked participants to indicate what had helped them to form their opinions about sharing study data with repositories. As with the previous question, participants were allowed to select all the responses that applied and a space was provided for the addition of responses that were not included in this question. "Conversations with my colleagues" was selected 24 times and "publications/conference sessions about sharing TBI data" was selected 13 times. "Publications/conference

sessions about TBI data sharing initiatives or established repositories" and "conversations with data curators" were both selected nine times. "Conversations with my supervisors" was selected eight times and "I did not form my opinions with the help of outside sources" was selected twice. Two participants who provided additional responses indicated that "online conversations and blogs" and "social media" had also helped them to form their opinions about sharing study data with repositories. Since these participants provided very similar answers of their own to the same question, it is possible that a response such as "conversations over social media channels" should have been included with the others. Doing so might have improved the original question's range of answer choices and increased the validity of the participants' answers.

Unfortunately for data curators, what is more probable is that TBI researchers are relying on their colleagues to help them form their opinions about sharing data with repositories rather than relying on the professionals who actually work for those repositories. However, data curators may be exerting more of an influence on these researchers than their answers suggest. Since these researchers appear to collect a fair amount of information on data sharing from publications and conference sessions, as well, it is possible that data curators were actually responsible for writing and presenting some of them. In retrospect, this detail might have been discovered if an additional survey question had asked respondents to specify whether the conference sessions they attended and the publications they read were contributed by data curators or other TBI researchers.

Arguments in favor of data sharing

The survey's next question asked participants to specify which arguments in favor of data sharing they had encountered through the conversations, publications, conference sessions, and other means they selected in the previous question. Once again,

participants were allowed to select all the responses that applied and a space was provided for the addition of responses that were not included in this question. “Data sharing promotes scientific transparency by making studies reproducible and their datasets verifiable” was the response that was selected most frequently (24 times), followed closely by “new knowledge often comes from the secondary analysis of shared datasets,” which was selected 23 times. The response “data sharing helps researchers avoid repeating studies unnecessarily” was selected 21 times, also followed closely by “data sharing increases the utility of data provided by human research subjects,” which was selected 20 times. The responses “data should be shared when the studies are paid for with public funds” and “research is translated into clinical practice more quickly when data is shared” were both selected 14 times. Finally, the response “datasets that are shared with repositories are better protected from loss and technological obsolescence” was selected nine times, while “I have not encountered any arguments in favor of research data sharing” was selected only once.

When asked if they agreed with the arguments in favor of research data sharing they had encountered, 19 participants answered “most of them.” Six participants answered “all of them” and three participants answered “some of them.” The question’s lone response of “not applicable” came from the researcher who answered, “I have not encountered any arguments in favor of research data sharing” to the previous question. None of the participants indicated that they did not agree with any of the arguments they had encountered.

As the reader may recall, two of the research questions that guide this study ask, “Are TBI researchers aware of the prevailing reasons for sharing research data? Do the TBI researchers who are aware of these reasons agree with them?” If one considers the participants’ answers to the question about which arguments in favor of data sharing they have encountered, then it is reasonable to suggest that most TBI researchers are

quite aware of the prevailing reasons. However, if one decides that most TBI researchers agree with the reasons they have encountered based solely on the participants' answers to the corresponding question, that conclusion would be incomplete without also considering the participants' explanations for these answers.

For example, one of the researchers who agreed with most of the arguments they had encountered in favor of data sharing cautioned, "Data repositories are only as good as the commitment to maintaining them and making them usable to researchers." Another researcher who also agreed with most of the arguments they had encountered in favor of data sharing pointed out that "in certain cases, the [intellectual property] rights may come into play." Yet the most incongruous explanation came from a researcher who agreed with all of the arguments they had encountered in favor of data sharing but believed it was only possible in "an ideal world." This researcher wrote, "Every laboratory uses a different approach to study design, data collection, and analysis. . . . There is no way to harmonize the data between labs without significant effort and money." If one examines these pieces of information together as a whole, then it may be more accurate to conclude that most TBI researchers agree with the reasons for data sharing in principle but are also quite aware of the issues to be resolved before data sharing becomes standard practice.

Experiences sharing study data with repositories

The survey's next question asked participants if they had ever submitted data from one of their own studies to a repository. The responses to this question were almost split down the middle, as fifteen participants answered "yes" while fourteen participants answered "no." This also served as the survey's only branching question, so that the participants who answered "yes" were presented with Questions #10 and #11, while the participants who answered "no" were presented with Questions #12, #13, and #14.

Question #10 asked participants why they submitted their study data to a repository. Once again, participants were allowed to select all the responses that applied and a space was provided for the addition of responses that were not included in this question. “I wanted my study data to remain available to TBI researchers over time” was selected seven times and “I share my study data whenever possible” was selected six times. “I was required to do so by the organization that funded my study” was selected five times and two different responses were both selected four times: “I would receive professional recognition for sharing my study data” and “I was required to do so by the publisher of my manuscript.” Only two participants selected “I was required to do so by my employer.” Three participants also provided additional responses; two of these revealed that they had submitted their study data to a repository “to see if analysts can validate or extend our findings,” and because of an “agreement with [a] consortium.” The third response revealed that the participant had shared data from a study that did not focus on traumatic brain injuries. (This researcher is again the one who disclosed that their employer limits how much data they can share.)

Given the fact that entities such as funders and publishers have only mandated the sharing of study data for a short period of time, those who support the practice may find it encouraging that several participants indicated they would be given professional recognition for sharing their data. Since the promise of receiving this benefit informed part of these researchers’ decisions to share their data, this may signify a change in the traditional systems of recognizing and rewarding scientists for their work. For example, the traditional system of the academic world is mainly based on researchers giving scientific presentations and publishing study results (Ohmann et al., 2017). However, if these systems have broadened to include sharing study data – and perhaps reusing shared data in secondary analyses – then more TBI researchers may be able to take advantage of these new opportunities for professional advancement.

Question #11 asked participants why they chose a particular repository as the site for their study data. Once again, participants were allowed to select all the responses that applied and a space was provided for the addition of responses that were not included in this question. “My colleagues recommended that I use the repository” and “the organization that funded my study recommended/required that I use the repository” were both selected six times. “The helpfulness of the repository’s staff and/or documentation made it easier for me to submit my data there” was selected three times. “I wanted to submit my data to a repository that only collects TBI or neuroscience study data” and “my employer recommended/required that I use the repository” were both selected twice. Three different responses were selected only once: “a data curator recommended that I use the repository,” “the publisher of my manuscript recommended/required that I use the repository,” and “it did not take much time and effort to comply with the repository’s file format and supporting documentation requirements.” Few participants provided their own responses to this question but one explained that in addition to demanding less time and effort to comply with its file format and supporting documentation requirements, their chosen repository was “free and supported by my university.”

Question #12 asked participants why they have not submitted data from one of their own studies to a repository. Once again, participants were allowed to select all the responses that applied and a space was provided for the addition of responses that were not included in this question. It is worth noting that in this case, many of the participants’ responses were clustered together and all of them followed closely in frequency to the others. Four different responses were selected five times: “I did not receive informed consent from my study participants to share their data,” “I am concerned that other researchers will misinterpret my data,” “it would take too much time and effort to comply with a repository’s file format and supporting documentation

requirements," and "the repositories I have encountered do not provide staff assistance or helpful data submission guidelines." The response "I am concerned that other researchers will publish misleading secondary analyses based on my data" was selected four times and the response "my employer/funder/publisher does not permit me to share my data" was selected three times. Three different responses were selected twice: "I am concerned that other researchers will publish secondary analyses based on my data before I get the chance," "I am not comfortable sharing data that I spent so much time and effort collecting," and "I am not sure that other researchers would be interested in my data." The response "I am concerned that I will not be able to properly de-identify my data for publication" was selected only once and none of the participants selected "TBI researchers are not recognized when they publish study data." It is also worth noting that two of the participants who had not shared data from one of their own studies were currently in the process of doing so; the first researcher explained that they are "in [the] process of sharing data of [an] ongoing study" and the second added that the "data sharing agreement with my institution is pending." While it is unfortunate that this survey did not come at a time when these researchers' experiences with data sharing could be examined, those who support the practice may find it encouraging that an increasing number of TBI researchers appear to be sharing their study data.

However, what is truly unfortunate is that most of the participants who had not shared data from one of their own studies declined to do so partly because of the administration and policies of the repositories themselves. In fact, the responses that underscored their frustrations with repositories' file format and supporting documentation requirements and data submission processes edged out the responses that expressed their fears about other researchers misinterpreting their data and publishing misleading analyses based on their data. Since TBI researchers may be relying more on their colleagues to help them form their opinions about data sharing – which

was suggested via the responses to a previous survey question - they may also be discussing these frustrations with each other and causing some TBI researchers to reconsider sharing their data with repositories.

Question #13 asked participants if they believed it was possible to resolve the concerns they selected in Question #12 so that they would be willing to share their study data with a repository in the future. Of the 14 participants who had not shared data from one of their own studies, 11 answered "yes," three answered "not sure," and none of them answered "no." When asked in Question #14 to briefly explain their answer to Question #13, only three participants gave a response. (Regrettably, none of the participants who answered "not sure" to Question #13 gave an explanation for this response; it would have been useful to know why they were not completely sure if their concerns about data sharing could be resolved.) The first participant, who answered "yes" to Question #13, restated that the data sharing agreement with their institution is pending. The second participant, who also answered "yes" to Question #13, revealed that they worked for a government institution where data sharing is not allowed but did not actually explain how they thought this issue with their employer might be resolved. As for the third participant - who also answered "yes" to Question #13 and expressed frustration with the administration and policies of repositories in Question #12 – this researcher simply wrote, "If their [sic] was funding to do this and the data repositories made it less onerous, I would be willing to submit data."

The survey's next question was presented to all participants and asked if they had ever been required to submit data from one of their own studies to a repository but had no such requirements for another study. If participants had ever been in this situation, they were asked to compare the two experiences. If not, they were asked to select the response "not applicable" and leave the rest of the question blank. As nineteen participants selected "not applicable" and three participants declined to answer the

question, it is reasonable to suggest that few TBI researchers have been exposed to both experiences. Of the seven participants who had been in this situation, four of them stated explicitly that it took additional time and effort to share data from TBI studies. Another participant implied that this extra time and effort dissuaded them from sharing data whenever it was not compulsory by stating, “When [data sharing] was not required, I didn’t do it.” However, two of the researchers believed that these additional costs were exceeded by the benefits provided by sharing study data. As one of the researchers put it, “It is obviously easier to be spared the burden of submitting data to an external repository, but I believe that the benefits outweigh the inconvenience.”

Final thoughts on sharing study data

The survey’s final question asked participants to record any additional thoughts they had about sharing data from TBI research. (Again, readers should be aware that not all of the participants answered this question.) Several of the responses offered the participants’ thoughts on data sharing generally - such as the researcher who proposed that data contributors should be offered authorship credits on publications that used their data - but the majority of the responses addressed issues that were specific to sharing TBI study data. For example, one researcher wrote that they considered data sharing to be critical to the study of traumatic brain injuries, “particularly for groups where there are more limited investigators and patient samples (e.g., moderate to severe pediatric TBI, intimate partner violence, etc.).” Another researcher added that they believed international metadata standards should be applied to magnetic resonance imaging data, which would allow researchers to determine if repositories collected pre-transformed raw images, transformed images, or both. “Without some [metadata] standardization,” the researcher concluded, “the images would be useless to most people.”

Finally, many references have been made throughout this paper to the real and perceived costs associated with preparing study data for sharing with repositories. Yet the most extreme example of the actual costs incurred by some TBI researchers may be the one from the researcher who revealed:

> *I recently wrote a grant that required [the Federal Interagency Traumatic Brain Injury Research Informatics System] and was astounded at the amount of time – would need to hire a data manager at about 50% just to comply with this requirement. This is simply not feasible for a regular budget. They also wanted grantees to collect biological samples and submit them to another database from which samples would then be sent for analysis. This required too much in terms of sample volumes for young kids and increased costs including the fee for using the repository and all of the mailing costs for overnight mail (to and from repository with frozen biological samples).*

Conclusion

Barbara Wildemuth describes an exploratory study as one that seeks to define phenomena that are worth studying further (2009). Since this study was always meant to be exploratory in nature, I examined all survey responses more than once in order to identify other phenomena of interest. By the time I completed the data analysis phase of this study, I had identified three such phenomena: the costs associated with sharing TBI study data, the data submission requirements of repositories, and the reuse of shared TBI study data.

Many of the researchers who participated in this study indicated that the extra amounts of time, effort, and money needed to prepare TBI study data for sharing with a repository only complicates the process. At least one researcher even implied that these added complications deterred them from sharing TBI study data when it was not required. When one considers the expensive and fragile forms of data that some TBI researchers work with, such as magnetic resonance images and biospecimens, the costs of preparing this data for sharing can soar even higher. Based on these survey responses, I believe that it would be useful for future studies to propose methods of reducing these expenditures. For example, one such study could compare the costs and benefits of requiring TBI researchers, data curators or managers, and other institutional staff (clerical staff, post-doctoral researchers, etc.) to prepare data for sharing and make recommendations based on the findings.

In the chapter that describes this study's methodology, I mentioned that I wish to enter the field of data curation and that I am a supporter of data sharing. For these reasons, I became concerned when a number of participants indicated that they found it

difficult to deal with the data submission processes and file format and supporting documentation requirements of repositories. Yet I believe that their responses present data curators with other useful avenues for research; namely, surveys of the data submission requirements of repositories that collect TBI study data. These surveys could also feature any repositories that found it necessary to streamline their data submission requirements and describe which requirements they simplified to make data sharing less difficult for everyone.

When the participants of this study were asked why they had not submitted any of their own data to a repository, one researcher explained that it was because "most of my work is done with data shared by others." Although I find this comment intriguing because it suggests that scientists within the TBI research community are already reusing shared data for verification purposes, secondary analyses, and other work, it cannot be used to gauge how widespread this practice has become. Therefore, I believe that it would be useful for other future studies to investigate how prevalent the reuse of shared data is among TBI researchers. Related studies could also examine the characteristics of TBI researchers who tend to reuse shared study data or uncover which types of shared data are reused most often (e.g., data from animal subjects, data from psychological tests, neuroimaging data, etc.). If these studies can demonstrate that the reuse of shared study data is common among these researchers, they could help to validate the existence of repositories that hold information which may be critical in the treatment of traumatic brain injuries.

In addition to exploring these topics in future studies, I believe there are actions that researchers, funders, repositories, and employers can take now to help data sharing become standard practice within the TBI research community. First, researchers may find that it is easier to secure funding for sharing study data with repositories by lobbying the grant-making institutions for these monies, rather than expecting their

colleagues to pay for access to their data. One major organization that has already provided years' worth of funds for some of this work is the U.S. National Institutes of Health (NIH). In March 2003, the NIH recognized that "it takes time and money to prepare data for sharing" and announced that applicants seeking $500,000 or more in direct costs in any year of their proposed research could request funds for data sharing and archiving (National Institutes of Health, 2003). If researchers band together and use their combined resources to put a greater amount of pressure on grant-making institutions – especially those that require researchers to share data from any funded study – they may discover that they can negotiate even better deals than the one offered by the NIH.

Repositories might also help to increase data sharing within the TBI research community by reviewing their data submission guidelines regularly and editing them for clarity, removing any outdated procedures, and simplifying procedures that are overly complex. Repositories could also update their lists of accepted data types and file formats during these review periods, as well as provide researchers with the URLs for other trusted repositories that accept the data types and file formats that they cannot. Along similar lines, repositories could revise their websites so that they direct researchers to staff members who can help with specific issues, rather than supplying a generic email address for all queries (e.g., "help_desk@repositoryname.org"). If repositories made even one or two of these changes, not only could those actions decrease the workload for data curators (since it is usually easier to update any documentation when done regularly), they could also reduce the burden on researchers in terms of time and effort spent on complying with repositories' data submission requirements.

Finally, the organizations that employ TBI researchers might also help to increase data sharing by investing both financial and human capital in the work that is needed to prepare TBI data for deposit with repositories. Since budgets and staff sizes vary

significantly among these organizations, it may be useful for researchers to first describe what is absolutely necessary to complete this work and then ask administrators to specify what types of resources they are able to commit. In the case of larger organizations, this could mean hiring data curators or managers to fill temporary or even permanent positions within the TBI researchers' schools, departments, or research groups. Smaller organizations, however, may have to resolve parts of this issue more creatively. For example, they could hire students to perform the lengthy data-entry tasks that go into preparing data for sharing or provide overtime salary payments to clerical staff to do this work. Whatever the resources each employer can provide, they all add up to an investment in transforming data sharing from a laborious process into standard practice.

References

Allen Institute for Brain Science. (2014, May 12). Understanding traumatic brain injury. Retrieved from https://alleninstitute.org/what-we-do/brain-science/news-press/articles/understanding-traumatic-brain-injury

Department of Defense. (2015, February). *Plan to establish public access to the results of federally funded research* (United States, Department of Defense). Retrieved from http://www.dtic.mil/dtic/pdf/dod_public_access_plan_ feb2015.pdf

Digital Curation Centre. (2017). Why preserve digital data? Retrieved from http://www.dcc.ac.uk/digital-curation/why-preserve-digital-data

Exec. Order No. 13625, 3 C.F.R. (2012). Title: Improving access to mental health services for veterans, service members, and military families

Federal Interagency Traumatic Brain Injury Research Informatics System. (2014, March 27). FITBIR data sharing policy. Retrieved from https://fitbir.nih.gov/sites/default/files/assets/FITBIR_Data_Sharing_Policy.pdf

Federer, L. M., Lu, Y., Joubert, D. J., Welsh, J., & Brandys, B. (2015). Biomedical data sharing and reuse: Attitudes and practices of clinical and scientific research staff. *PLOS One,* 10(6). doi: 10.1371/journal.pone.0129506

International Consortium of Investigators for Fairness in Trial Data Sharing. (2016, August 4). Toward fairness in data sharing. Retrieved from http://www.nejm.org/doi/full/10.1056/NEJMp1605654

International Initiative for Traumatic Brain Injury Research. (2014). InTBIR mission. Retrieved from https://intbir.nih.gov/mission

International Initiative for Traumatic Brain Injury Research. (2016a, December 12). InTBIR "deliverables" survey summary. Retrieved from https://intbir.nih.gov/sites/all/themes/nivea//InTBIR%20deliverables%20survey%20summary.pdf

International Initiative for Traumatic Brain Injury Research. (2016b). InTBIR data sharing survey. Retrieved from https://intbir.nih.gov/sites/all/themes/nivea//Data_sharing_Survey_InTBIR_2016.pdf

Interview on data curation profile modules 5-13 [Personal interview]. (2017, March 27).

LeClere, F. (2010, August 03). Too many researchers are reluctant to share their data. Retrieved from http://www.chronicle.com/article/Too-Many-Researchers-Are/123749

Murray, T. (2016). Researchers oppose data-sharing proposal. *Canadian Medical Association Journal,* 188(14). doi:10.1503/cmaj.109-5317

National Institutes of Health. (2003, March 5). NIH data sharing policy and implementation guidance. Retrieved from https://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm

National Science Foundation. (2017, January 30). *Proposal & award policies & procedures guide* (United States, National Science Foundation). Retrieved from https://www.nsf.gov/pubs/policydocs/pappg17_1/pappg_11.jsp#XID4

O'Dowd, A. (2014). New incentives are needed to boost research data sharing, says expert group. *British Medical Journal,* 348(7961).

Ogburn, J. L. (2010). The imperative for data curation. *Portal: Libraries and the Academy,* 10(2), 241-246. doi:10.1353/pla.0.0100

Ohmann, C., Banzi, R., Canham, S., Battaglia, S., Matei, M., Ariyo, C., . . . Demotes-Mainard, J. (2017). Sharing and reuse of individual participant data from clinical trials: Principles and recommendations. *BMJ Open,* 7, 1-24.

doi:10.1136/bmjopen-2017-018647

Oliver, G., & Harvey, R. (2016). *Digital curation* (2nd ed.). Chicago, IL: Neal-Schuman.

TED Initiative. (2015, September 4). TBI endpoints development initiative research collaboration policy. Retrieved from http://tbiendpoints.ucsf.edu/sites/tbiendpoints.ucsf.edu/files/TED%20Research%20Collaboration%20Policy_9-4-2015.pdf

Tosetti, P., Hicks, R. R., Theriault, E., Phillips, A., Koroshetz, W., & Draghia-Akli, R. (2013). Toward an international initiative for traumatic brain injury research. *Journal of Neurotrauma,* 30(14), 1211-1222. doi:10.1089/neu.2013.2896

Traumatic Brain Injury Model Systems National Data and Statistical Center. (2015, June 12). External use TBIMS National Database notification. Retrieved from https://www.tbindsc.org/StaticFiles/SOP/602d%20-%20External%20Use%20TBIMS%20National%20Database%20Notification.pdf

U.S. Department of Veterans Affairs. (2017, August 21). VA research on traumatic brain injury (TBI). Retrieved from https://www.research.va.gov/topics/tbi.cfm#research1

Weightman, M. M., Radomski, M. V., Finkelstein, M., Cecchini, A. S., Davidson, L. F., Heaton, K. J., . . . Scherer, M. R. (2017). Further development of the Assessment of Military Multitasking Performance: Iterative reliability testing. *PLOS One,* 12(1). doi:10.1371/journal.pone.0169104

Wildemuth, B. M. (2009). *Applications of social research methods to questions in information and library science.* Westport, CT: Libraries Unlimited.

Appendix A: Initial Recruitment Email

Dear Researcher,

My name is Nicole Robertson and I am an Information Science Master's student at the University of North Carolina at Chapel Hill. I am currently conducting research for a Master's Paper on the attitudes of traumatic brain injury (TBI) researchers toward sharing study data with repositories.

The purpose of this study is to collect information on TBI researchers' perceptions of data sharing and their willingness to share study data via repositories. This study seeks to identify the reasons why scientists within this population are or are not in favor of sharing data via repositories. In addition, this study seeks to identify which of the prevailing reasons in favor of sharing data these scientists are aware of and if they agree with them. Doing so will give both the TBI research community and data curators a better understanding of the variety of opinions on data sharing that exist within this population.

You are invited to take part in this study by completing an online survey. The link to the survey is provided below. Your participation in this study is completely voluntary and will take about 20 minutes. Please feel free to forward the survey link to other TBI researchers who may be interested in giving their opinions on data sharing.

[Link to Qualtrics version of survey]

The data resulting from this study may be published in a repository (e.g., UNC Dataverse). However, the survey will not collect any personally identifying information and survey responses will be recorded anonymously so that no one can link your answers to you personally.

Thank you for your time and consideration. If you have any questions about this study or its procedures, please contact me via email, [email address].

Sincerely,

Nicole Robertson

MSIS Candidate 2019, UNC-Chapel Hill

Appendix B: Survey Questions

1. Please select the most appropriate answer: Traumatic brain injury (TBI) researchers should share their study data with repositories.

□ All of the time □ Most of the time □ Some of the time □ Seldom □ Never

2. Please briefly explain your answer.

3. I would be concerned if _____ expected me to share my study data with a repository. (Please select all that apply.)

□ My employer □ My colleagues □ A publisher that accepted one of my manuscripts □ The organization that provided me with a research grant □ Other (Please specify) □ This expectation would not concern me

4. Please briefly explain your answer.

5. _____ helped me to form my opinions about sharing study data with repositories. (Please select all that apply.)

□ Conversations with my colleagues □ Conversations with my supervisors □ Conversations with data curators (professionals who make datasets within repositories discoverable, accessible, and usable over time) □ Publications/conference sessions about sharing TBI data □ Publications/conference sessions about TBI data sharing initiatives or established repositories □ Other (Please specify) □ I did not form my opinions with the help of outside sources

6. Which arguments in favor of research data sharing have you encountered through the conversations, publications, conference sessions, and other means you selected in the previous question? (Please select all that apply.)

□ New knowledge often comes from the secondary analysis of shared datasets □ Research is translated into clinical practice more quickly when data is shared □ Data sharing promotes scientific transparency by making studies reproducible and their data verifiable □ Datasets that are shared with repositories are better protected from loss and technological obsolescence □ Data should be shared when the studies are paid for with public funds □ Data sharing helps researchers avoid repeating studies unnecessarily □ Data sharing increases the utility of data provided by human research subjects □ Other (Please specify) □ I have not encountered any arguments in favor of research data sharing

7. Do you agree with the arguments you have encountered?

□ All of them □ Most of them □ Some of them □ None of them □ Not applicable

8. Please briefly explain your answer.

9. Have you ever submitted data from one of your own studies to a repository?

□ Yes □ No

(**BIFURCATED QUESTION:** If participants answer "Yes," they will be presented with Questions #10 and #11; if participants answer "No," they will be presented with Questions #12, #13, and #14.)

10. Why did you submit your study data to a repository? (Please select all that apply.)

□ I was required to do so by my employer □ I share my study data whenever possible □ I would receive professional recognition for sharing my study data □ I was required to do so by the publisher of my manuscript □ I wanted my study data to remain available to TBI researchers over time □ I was required to do so by the organization that funded my study □ Other (Please specify)

11. Why did you choose that particular repository as the site for your study data? (Please select all that apply.)

□ I wanted to submit my data to a repository that only collects TBI or neuroscience study data □ A data curator recommended that I use the repository □ My colleagues recommended that I use the repository □ The helpfulness of the repository's staff and/or documentation made it easier for me to submit my data there □ The organization that funded my study recommended/required that I use the repository □ The publisher of my manuscript recommended/required that I use the repository □ My employer recommended/required that I use the repository □ It did not take much time and effort to comply with the repository's file format and supporting documentation requirements □ Other (Please specify)

12. Why have you not submitted data from one of your own studies to a repository? (Please select all that apply.)

□ I am concerned that other researchers will publish secondary analyses based on my data before I get the chance □ I am concerned that I will not be able to properly de-identify my data for publication □ I did not receive informed consent from my study participants to share their data □ I am concerned that other researchers will misinterpret my data □ I am concerned that other researchers will publish misleading secondary analyses based on my data □ It would take too much time and effort to comply with a repository's file format and supporting documentation requirements □ The repositories I have encountered do not provide staff assistance or helpful data submission guidelines □ TBI researchers are not recognized when they publish study data □ I am not sure that other researchers would be interested in my data □ My employer/funder/publisher does not permit me to share my data □ I am not comfortable sharing data that I spent so much time and effort collecting □ Other (Please specify)

13. Do you believe it is possible to resolve the concerns you selected in the previous question so that you would be willing to share your study data with a repository in the future?

□ Yes □ No □ Not sure

14. Please briefly explain your answer.

15. Have you ever been required to submit data from one of your own studies to a repository but had no such requirements for another study? How would you compare the two experiences?

□ Not applicable, I have never been in this situation

16. Please use this space to record any additional thoughts you may have about sharing data from TBI research.